PCT/GB 00/03860

INVESTOR IN PEOPLE

GB 00/3860

4

10/089877

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

REC'D 27 OCT 2000
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

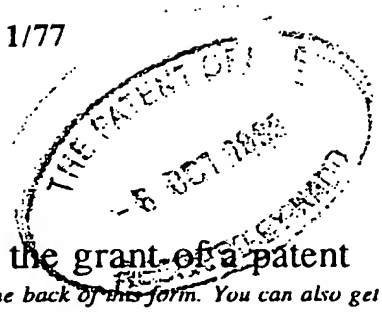
Dated 12/10/00

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN

Patents Form 1/77

Patents Act 1977
(Rule 16)



The Patent Office

0705199 E493167-4 D02890
P01/7700 0.00 - 9923644.0

Request for the grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — REP06193GB

2. Patent *(The P.* — **9923644.0**

-6 OCT 1999

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Medical BioSystems Ltd.
The Old Mill
Beaston Cross
Broadhempston
Nr. Totnes
Devon TQ9 6BX
United Kingdom

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation — GB

4. Title of the invention — DNA SEQUENCING

5. Name of your agent *(if you have one)* — GILL JENNINGS & EVERY

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Broadgate House
7 Eldon Street
London
EC2M 7LH

Patents ADP number *(if you know it)* — 745002

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

NO

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	8
Claim(s)	2
Abstract	
Drawing(s)	

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. For the Applicant
Gill Jennings & Every

I/We request the grant of a patent on the basis of this application.

Signature [signature] Date 6 October 1999

12. Name and daytime telephone number of person to contact in the United Kingdom

HUCKER, Charlotte Jane
0171 377 1377

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

DNA SEQUENCING

Field of the Invention

This invention relates to polynucleotide sequence determinations.

Background of the invention

The ability to determine the sequence of a polynucleotide is of great scientific importance. For example, the Human Genome Project is an ambitious international effort to map and sequence the three billion bases of DNA encoded in the human genome. When complete, the resulting sequence database will be a tool of unparalleled power for biomedical research. The major obstacle to the successful completion of this project concerns the technology used in the sequencing process.

The principle method in general use for large-scale DNA sequencing is the chain termination method. This method was first developed by Sanger and Coulson (Sanger et al. Proc. Natl. Acad. Sci. USA 1977; 74: 5463-5467), and relies on the use of dideoxy derivatives of the four nucleoside triphosphates which are incorporated into the nascent polynucleotide chain in a polymerase reaction. Upon incorporation, the dideoxy derivatives terminate the polymerase reaction and the products are then separated by gel electrophoresis and analysed to reveal the position at which the particular dideoxy derivative was incorporated into the chain.

Although this method is widely used and produces
ble results, it is recognised that it is slow, labour-
and expensive.

ternative sequencing method is proposed in EP-A-
ch uses spectroscopic means to detect the
a nucleotide into a nascent polynucleotide
y to a target. The method relies on an
of template and primer, which is
ning only one of the different
techniques are then used to
rising from the polymerase

enzymatic cross-over between

catalysed growth of the template copy. The spectroscopic techniques described are surface plasmon resonance (SPR) spectroscopy, which measures changes in an analyte within an evanescent wave field, and fluorescence measuring techniques. However. limitations of this method are recognised; the most serious for the SPR technique being that, as the size of the copy strand grows, the absolute size of the signal also grows due to the movement of the strand out of the evanescent wave field, making it harder to detect increments. The fluorescence measuring techniques have the disadvantage of increasing background interference from the fluorophores incorporated on the growing nascent polynucleotide chain. As the chain grows, the background "noise" increases and the time required to detect each nucleotide incorporation needs to be increased. This severely restricts the use of the method for sequencing large polynucleotides.

Single fragment polynucleotide sequencing approaches are outlined in WO-A-9924797 and WO-A-9833939, both of which employ fluorescent detection of single labelled nucleotide molecules. These single nucleotides are cleaved from the template polynucleotide, held in a flow by an optical trap (Jett, J.H. *et al*. J. Biomol. Struc. Dyn (1989) 7:301-309, by the action of an exonuclease molecule. These cleaved nucleotides then flow downstream within a quartz flow cell, are subjected to laser excitation and then detected by a sensitive detection system. However, limitations of this method are recognised; the most serious for the exonuclease technique being the fact that the labelled nucleotides severely affect the processivity of the exonuclease enzyme. Other limitations of this method include 'sticking' of the nucleotide(s) to the biotin bead used to immobilise the polynucleotide fragment thus resulting in the nucleotide flow becoming out of phase, inefficiency and length limitation of the initial enzymatic labelling process; and the excitation 'cross-over' between

the four different dye molecules resulting in a greatly increased error rate.

There is therefore a need for an improved method, preferably at the single fragment level, for determining the sequence of polynucleotides which significantly increases the rate and fragment size of polynucleotide sequenced and which is preferably carried out by an automated process, reducing the complexity and cost associated with existing methods.

Summary of the Invention

The present invention is based on the realisation that the sequence of a target polynucleotide can be determined by measuring conformational changes in an enzyme that binds to and processes along the target polynucleotide. The extent of the conformational change that takes place is different depending on which individual nucleotide on the target is in contact with the enzyme.

According to one aspect of the present invention, a method for determining the sequence of a polynucleotide comprises the steps of:

(i) reacting a target polynucleotide with an enzyme that is capable of interacting with and processing along the polynucleotide, under conditions sufficient to induce the enzyme activity; and

(ii) detecting conformational changes in the enzyme as the enzyme processes along the polynucleotide.

According to one embodiment of the invention, the enzyme is a polymerase enzyme which interacts with the target in the process of extending a complementary strand. The enzyme is typically immobilised on a solid support to localise the reaction within a defined area.

According to a second embodiment of the invention, the enzyme comprises a first bound detectable label, the characteristics of which alter as the enzyme undergoes a conformational change. The enzyme may also comprise a [illegible]able label capable of interacting with [illegible] the degree of interaction is

dependent on a conformational change in the enzyme. Typically, the first label is an energy acceptor and the second label is an energy donor, and detecting the conformational change is carried out by measuring energy transfer between the two labels.

According to a further embodiment of the invention, fluorescence resonance energy transfer (FRET) is used to detect a conformational change in an enzyme that interacts with and processes along a target polymerase, thereby determining the sequence of the polynucleotide. Fluorescence resonance energy transfer may be carried out between FRET donor and acceptor labels, each bound to the enzyme. Alternatively, one of the labels may be bound to the enzyme and the other label bound to the polynucleotide.

According to a further embodiment, there is the use of a detectably-labelled enzyme, capable of interacting with and processing along a target polynucleotide, to determine the sequence of the polynucleotide, wherein the label alters its detectable characteristics as the enzyme processes along the polynucleotide.

The present invention offers several advantages over conventional sequencing technology. Once a polymerase enzyme begins its round of polynucleotide elongation, it tends to polymerase several thousand nucleotides before falling off from the strand. Additionally, certain specific polymerase systems are able to anchor or tether themselves to the template polynucleotide via a 'sliding clamp' (e.g. Polymerase III) which encircles the template molecule or via a molecular hook (e.g. T7:thireodoxin complex) which partially encircles the template.

The invention may also enable tens of kilobases (kb) or more to be sequenced in one go at a rate of hundreds of base pairs per second. This is a result of the fact that the sequencing process is carried out on a single fragment of DNA. An advantage of sequencing a single fragment of DNA is that sequencing rates are determined by the enzyme system utilised and not upon indirect, summated reactions,

and are therefore correspondingly higher. Just as important as the high rate is the ability to sequence large fragments of DNA. This will significantly reduce the amount of subcloning and the number of overlapping sequences required to assemble megabase segments of sequencing information. An additional advantage of single fragment approaches is the elimination of problems associated with the disposal of hazardous wastes, such as acrylamide, which plague current sequencing efforts.

Description of the Invention

The present method for sequencing a polynucleotide involves the analysis of conformational changes between a enzyme and a target polynucleotide.

The term "polynucleotide" as used herein is to be interpreted broadly, and includes DNA and RNA, including modified DNA and RNA, as well as other hybridising nucleic acid-like molecules, e.g. peptide nucleic acid (PNA).

The enzyme may be a polymerase enzyme, and a conformational change is brought about when the polymerase incorporates a nucleotide into a nascent strand complementary to the target polynucleotide. It is expected that the conformational change will be different for each of the different nucleotides, A, T, G or C and therefore measuring the change will identify which nucleotide is incorporated.

Alternatively, the enzyme may be any that is involved in an interaction with a polynucleotide, e.g. a helicase enzyme, primase and holoenzymes. As the enzyme processes along the polynucleotide, its conformation will change depending on which nucleotide on the target it is brought into contact with.

One way of detecting a conformational change in the enzyme is to measure resonance energy transfer between a suitable energy donor label and a suitable energy acceptor label. In one example the donor and acceptor are each bound to the enzyme and the conformational change in the enzyme brought about by its interaction with the target

polynucleotide alters the relative positioning of the labels. The differences in positioning are reflected in the resulting energy transfer and are characteristic of the particular nucleotide in contact with the enzyme. Alternatively, one label may be positioned on the enzyme and the other on a nucleotide of the target.

The use of fluorescence resonance energy transfer (FRET) is a preferred embodiment of the invention. This technique is capable of measuring distances on the 2- to 8nm scale and relies on the distance-dependent energy transfer between a donor fluorophore and an acceptor fluorophore. The technique not only has superior static co-localization capabilities but can also provide information on dynamic changes in the distance or orientation between the two fluorophores for intramolecular and intermolecular FRET. Since the first measurement of energy transfer between a single donor and a single acceptor (single pair FRET) (Ha, T. *et al* (1996) Proc. Natl. Acad. Sci. USA 96:893), it has been used to study ligand-receptor co-localisation (Schutz, G. J. *et al.* (1998) Biophys. J. 74:2223), to probe equilibrium protein structural fluctuations and enzyme-substrate interactions during catalysis (Ha, T. *et al* (1999) Proc. Natl. Acad. Sci. USA 96:893), and to identify conformational states and sub-populations of individual diffusing molecules in solutions. All of these variables are envisioned as applicable within the context of the invention.

The present invention may also be carried out using measurement techniques that require only a single label. Any system that is capable of measuring changes in the local environment of the enzyme at the single molecule level, is an accepted embodiment of the invention. Various properties of single fluorescent probes attached to a polynucleotide processive enzyme and/or its substrate(s) can be exploited in the context of the invention to provide data on variables within or in close proximity to the enzyme system/molecular environment that are specific to a

nucleotide incorporation event. Such variables include molecular interactions, enzymatic activity, reaction kinetics, conformational dynamics, molecular freedom of motion, and alterations in activity and in chemical and electrostatic environment.

For example, the absorption and emission transition dipoles of single fluorophores can be determined by using polarized excitation light or by analysing the emission polarisation, or both. The temporal variation in dipole orientation of a rigidly attached or rotationally diffusing tethered label can report on the angular motion of a macromolecule system or one of its subunits (Warshaw, D. M. *et al* (1998) Proc. Natl. Acad. Sci. USA 95:8034) and therefore may be applied in the present invention.

The labels that may be used in the present invention will be apparent to those skilled in the art. Preferably, the label is a fluorescence label, such as those disclosed in Xue, Q. *et al.* (1995) Nature 373:681. Alternatively, fluorescing enzymes (Lu, H. P. *et al* (1998) Science 282:1877) can be employed.

The preferred embodiment of the invention, however, involves the use of small fluorescence molecules that are covalently and site-specifically attached to the polynucleotide processive enzyme, e.g. tetramethylrhodamine (TMR).

Another foreseen embodiment used to tag the polynucleotide processive enzyme is the fusion of green fluorescent protein (GFP) to the processive enzyme (e.g. polymerase) via molecular cloning techniques known in the art (Pierce, D.W. *et al* (1997) Nature 388:338). This technique has been demonstrated to be applicable to the measurement of conformational changes (Miyawaki, A *et al* (1997) Nature 388:882) and local pH (Llopis, J. *et al.* (1998) Proc. Natl. Acad. Sci. USA 95:6803).

A number of strategies may be used to attach the fluorescent dyes to the enzyme. Strategies include the use of site-directed mutagenesis and unnatural amino acid

mutagenesis (Anthony-Cahil, S.J. *et al* (1989) Trends Biochem. Sci. 14:400) to introduce cysteine and ketone handles for specific and orthogonal dye labeling proteins (Cornish, V.W. *et al*. (1994) Proc. Natl. Acad. Sci. USA 91:2910).

A variety of techniques may be used to measure a conformational change in the enzyme. Resonance energy transfer may be measured by the techniques of surface plasmon resonance (SPR) or fluorescent surface plasmon resonance may be used.

However, other techniques which measure changes in radiation via interaction with a 'label' or energy transducer may be considered, for example spectroscopy by total internal reflectance fluorescence (TIRF), attenuated total reflection (ATR), frustrated total reflection (FTR), Brewster angle, reflectometry, scattered total internal reflection (STIR) or evanescent wave ellipsometry.

An important aspect of the method of the present invention is the use of a single enzyme, appropriately labelled. Use of, for example, a single labelled polymerase offers several important advantages for the success of this method. Firstly, the problem of intermittent processivity of non-polymerase molecules (e.g. exonucleases) in single fragment environments is reduced considerably. Secondly, the problem of having to detect single labelled molecules (i.e. nucleotides) within a flow stream and its inherent noise problems is avoided. This also removes the problem of uncontrolled nucleotide binding to surfaces related to or within the template polynucleotide.

CLAIMS

1. A method for determining the sequence of a polynucleotide, comprising the steps of:

i. reacting a target polynucleotide with an enzyme that is capable of interacting with and processing along the polynucleotide, under conditions sufficient to induce enzyme activity; and

ii. detecting conformational changes in the enzyme as the enzyme processes along the polynucleotide.

2. A method according to claim 1, wherein the enzyme is a polymerase enzyme.

3. A method according to claim 1 or claim 2, wherein the enzyme is immobilised on a solid support.

4. A method according to any preceding claim, wherein the enzyme comprises a first bound detectable label, the characteristics of which alter as the enzyme undergoes a conformational change.

5. A method according to claim 4, wherein the enzyme comprises a second bound detectable label capable of interacting with the first label, wherein the degree of interaction is dependent on a conformational change in the enzyme.

6. A method according to claim 4, wherein a second detectable label is bound to a nucleotide brought into contact with the enzyme.

7. A method according to claim 5 or claim 6, wherein the first label is an energy acceptor and the second label is an energy donor, and wherein step (ii) is carried out by measuring energy transfer between the two labels.

8. A method according to claim 4, wherein step (ii) is carried out by measuring a polarisation effect consequent on the altered characteristics of the first label.

9. Use of fluorescence resonance energy transfer to detect a conformational change in an enzyme that interacts with and processes along a target polymerase and thereby determine the sequence of the polynucleotide.

10. Use according to claim 9, wherein the enzyme is a polymerase enzyme.

11. Use according to claim 9 or claim 10, wherein the enzyme is immobilised on a solid support.

12. Use of a detectably-labelled enzyme, capable of interacting with and processing along a target polynucleotide, to determine the sequence of the polynucleotide, wherein the label alters its detectable characteristics as the enzyme processes along the polynucleotide.